

Laboratoire de statistique

STT3781

Hiver 2025
Janie Coulombe
Tél. 514-343-7977

3 crédits
Bureau: 4243 Pavillon André Aisenstadt
janie.coulombe@umontreal.ca

Horaire des cours et ateliers

Du 9 janvier au 14 avril 2025. Il y a une pause pendant la semaine de relâche (semaine du 3 mars). Horaire:

- Lundi* de 10h30 à 12h20
- Jeudi de 11h30 à 12h20

*Sauf les 20 et 27 janvier, à discuter en classe (horaire sujet à changement).

Salle de cours

Attention: La salle de cours changera au cours de la session et il y a plusieurs incertitudes quant à l'endroit où les cours se tiendront. C'est sujet à changement. Pour le moment, nous considérerons que les cours du:

- Lundi sont au laboratoire, au 4191 Pav. Andre-Aisenstadt
- Jeudi sont au 1409 Pav. André-Aisenstadt

(comme montré sur votre centre étudiant).

Examen

Un examen final prévu pour le 17 avril, à emporter et à remettre le 28 avril (dates sujettes à changement). Il n'y aura pas d'examen final en présentiel le 23 malgré qu'il soit indiqué cela sur votre centre étudiant.

Disponibilité

Ma période de disponibilité sera chaque jeudi de 9h30 à 11h30 à mon bureau.

Note: Il n'y a pas de séance de travaux pratiques dans ce cours. Des cours magistraux ainsi que des cours de type atelier seront donnés en alternance. Vous présenterez aussi vos résultats de façon orale à quelques reprises.

Description

Dans ce cours, nous utilisons les connaissances théoriques et l'expérience acquises dans les cours précédents en statistique pour répondre à différentes questions de recherche. Cinq projets d'équipe seront présentés durant la session. Pour chaque projet, les étudiants seront appelés à analyser un jeu de données à partir d'un des trois logiciels statistiques discutés. On apprendra à examiner et décrire un jeu de données, discuter de détails importants dans l'analyse de données et trouver une méthode d'analyse appropriée, et interpréter et divulguer les résultats de façon claire et concise. En outre, les étudiants seront appelés à présenter les résultats de leurs analyses lors de présentations orales en équipe ou dans un rapport statistique.

Les projets (travaux en équipe) porteront sur 5 sujets parmi la visualisation de données, la régression logistique, l'inférence causale, les études de simulation pour le calcul de puissance, les données manquantes, l'apprentissage automatique et les mesures répétées (pas nécessairement dans cet ordre). L'examen final est un projet à emporter à la maison et consiste à décrire et analyser un jeu de données pour répondre à une question précise. Le projet nécessite l'écriture d'un rapport (ce n'est pas en équipe).

Objectifs du cours

- Vous exposer à des problèmes réels d'analyse de données
- Explorer et analyser les données à l'aide de logiciels (SAS, R, SPSS)
- Vous pratiquer à écrire des rapports scientifiques dans lesquels vous décrierez la problématique, les méthodes utilisées et les résultats obtenus

- Améliorer la communication avec vos pairs (et le client), et vous fournir des occasions de présenter vos résultats devant vos pairs et d'en discuter en groupe
- Apprendre sur certaines méthodes statistiques couramment utilisées

Évaluation

- 5 projets, valant au total 65% de la note finale au cours
- 10 points de participation: Plus d'informations seront données au premier cours, quant à la façon dont ces points seront distribués
- Un projet final valant pour 25% de la note finale au cours

Important: Il faut une note minimale de 50% à l'examen final pour passer le cours (en plus d'avoir une moyenne de plus de 50 de manière générale).

Important: Les étudiant.e.s qui ne se présenteront pas lors d'une présentation orale d'équipe (sans raison valable) perdront automatiquement leurs points pour le projet à présenter ce jour-là.

Autres

La date limite pour abandonner le cours sans frais est le 23 janvier 2025 et la date limite pour abandonner avec frais est le 14 mars 2025. Si vous abandonnez après cette date, vous aurez automatiquement un échec. De plus, il appartient à l'autorité compétente de décider si une absence au final est motivée ou non (vous devez mentionner cette absence le plus tôt possible).

Les étudiant.e.s inscrit.e.s au Bureau de Soutien aux Étudiants en Situation de Handicap (BSESH) désirant bénéficier de mesures d'accommodement aux examens sont priés de consulter le lien suivant pour connaître la procédure à suivre: <https://safire.umontreal.ca/reussite-et-ressources/mesures-daccommodement-aux-examens-pour-les-etudiants-en-situation-de-handicap/>.

Plagiat

L'Université de Montréal a une politique très claire en ce qui concerne le plagiat. Vous êtes invité.e.s à la consulter au www.integrite.umontreal.ca. Cette politique concerne non seulement les examens mais aussi les devoirs.

Important: Je porte votre attention sur la section décrivant les types d'infraction, qui comporte entre autres comme infraction l'utilisation de contenus générés par un système d'intelligence artificielle.

Autres ressources importantes:

- Si vous éprouvez des difficultés dans votre processus d'apprentissage, et que vous souhaitez recevoir des conseils, vous pouvez contacter Karima Amoura, conseillère à la réussite étudiante, à l'adresse suivante : cre@dms.umontreal.ca
- Le centre de santé et de consultation psychologique (CSCP) de l'Université de Montréal (<http://www.cscp.umontreal.ca/>) - la prise de rendez-vous et l'inscription à un premier rendez-vous se font entièrement en ligne.
- Le programme mieux-être de l'ASEQ (ligne téléphonique ouverte 24 heures/7 jours sur 7 au 1-833-851-1363) - plus d'informations au http://www.aseq.ca/rte/fr/FACUM_Programmedaide_Programmedaide
- N'hésitez pas à contacter votre TGDE (technicienne en gestion des dossiers étudiants) à tgdebac@dms.umontreal.ca ou votre association étudiante (aemsum@dms.umontreal.ca) qui pourront vous guider.

Bibliographie

Ouvrages pertinents:

- Régression logistique:
 - Agresti, Alan. Categorical data analysis. John Wiley & Sons, 2003.
 - Christensen, Ronald. Log-linear models and logistic regression. Springer Science & Business Media, 2006.
- Inférence causale:

- Arel-Bundock, Vincent. Analyse causale et méthodes quantitatives: Une introduction avec R, Stata et SPSS. Presses de l'Université de Montréal, 2020.
- Hernan, Miguel A. et Robins, James M. Causal inference: What If, 1ère édition. 2010. Se trouve ici gratuitement: <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>.
- Données manquantes et imputation:
 - Little, Roderick JA et Rubin, Donald B. Statistical analysis with missing data. John Wiley & Sons, 2019.
 - Van Buuren, Stef. Flexible imputation of missing data. CRC press, 2018.
- Apprentissage automatiques:
 - G James, D Witten, T Hastie et R Tibshirani: An introduction to statistical learning. Springer, 2013.
 - T Hastie, R Tibshirani et J. Friedman: The elements of statistical learning - Data mining, inference and prediction. Springer, 2009.
- Mesures répétées:
 - West, Brady T, Welch, Kathleen B, et Galecki, Andrzej T. Linear mixed models: A practical guide using statistical software. Chapman and Hall/CRC, 2006.
 - Verbeke, Geert, et Molenberghs, Geert. Linear mixed models for longitudinal data. New York: Springer, 2000.
- Programmation:
 - Lafaye de Micheaux, Pierre, Drouilhet, Rémy, et Liquet, Benoit. Le logiciel R-Maîtriser le langage-Effectuer des analyses statistiques. HAL, 2010, vol. 2010.

Pour la disponibilité des livres en bibliothèque, vous pouvez contacter le comptoir de prêt au www.bib.umontreal.ca/nous-joindre/mi.htm ou Mme Indiana Delsart, bibliothécaire, au indiana.delsart@umontreal.ca.